

Amino acid sequence of cinnamomin, a new member of the elicitin family, and its comparison to cryptogein and capsicein

Jean-Claude Huet and Jean-Claude Pernollet

Laboratoire d'Etude des Protéines, Centre INRA, route de St Cyr, 78026 Versailles Cedex, France

Received 5 September 1989

The phytopathogenic fungi *Phytophthora cinnamomi* cause systemic leaf necrosis on its non-host tobacco; in culture, it secretes a protein, called cinnamomin, which elicits leaf necrosis and protects tobacco against the pathogen *Phytophthora nicotianae*, in a way similar to cryptogein and different from capsicein, elicitors of known amino acid sequences. The cinnamomin sequence has been determined and compared to other elicitors. The differences in the 3 elicitor sequences were correlated to the biological activities: 2 lysines were ascribed as the key amino acids involved in the differential control of protection with respect to necrosis.

Amino acid sequence; Elicitor; Plant pathogen; (*Phytophthora*, *Nicotiana tabacum*)

1. INTRODUCTION

Upon inoculation, *Phytophthora nicotianae*, the agent of the tobacco black shank, invades tobacco stems, whereas other *Phytophthora* species cause limited colonization accompanied by leaf necrosis at a distance from the inoculation court. In contrast, no distant leaf necrosis is observed during the interaction between *P. nicotianae* and tobacco, and no elicitor of necrosis has been detected in culture filtrates [1]. In culture, *P. cryptogea*, *P. capsici* and *P. cinnamomi* secrete low molecular mass proteins named cryptogein, capsicein and cinnamomin, respectively [2], which have been purified [3]. These fungal proteins constitute a novel protein family, called elicitors. Applied on tobacco plants, elicitors not only elicit leaf necrosis, but also cause the accumulation of pathogenesis-related proteins [4] and induce protection against a subsequent inoculation with the tobacco pathogen *P. nicotianae*. Elicitors from various species exhibit different levels of biological activities, abolished upon pronase digestion [5]: cryptogein causes visible leaf necrosis when applied at ~1 µg per plant, whereas 50-fold as much capsicein is required for the same reaction; in contrast to cryptogein, capsicein induces protection even in the near absence of leaf necrosis. The complete amino acid sequences of these two elicitors are known: they are of

similar M_r (respectively 10 323 and 10 155) and consist of 98 residues. They are holoproteins devoid of glycosylation which exhibit a 85% homology: only 2 short terminal regions, which should be involved in the modulation of the biological activities which distinguish cryptogein from capsicein, are heterologous [5].

Culture filtrates of *P. cinnamomi* induce protection of tobacco against its pathogen [4]. Cinnamomin, purified from these filtrates [3], shows a level of necrotic activity comparable to cryptogein [1,3] and protects tobacco against its pathogen at similar doses. We have therefore determined the amino acid sequence of cinnamomin to get a more precise knowledge of the amino acid residues which are involved in the modulation of the elicitor biological activities.

2. MATERIALS AND METHODS

2.1. Material

Cinnamomin was purified as already described [3] and was reduced with dithiothreitol and S-carboxymethylated with unlabeled iodoacetic acid according to Fernandez-Luna et al. [6].

2.2. CNBr, Asp-N and tryptic digests

CNBr peptides were obtained as follows: S-carboxymethylated protein was dissolved in 70% (v/v) formic acid (20 mg/ml). After addition of a 25-fold weight excess of cyanogen bromide over the methionine content, the mixture was incubated for 24 h at room temperature in the dark; the reaction mixture was then diluted 5-fold with cold distilled water and lyophilized. Asp-N endoproteinase (EC 3.4.21), which allows the cleavage at the amino acid side of Asp and Cys(SO₃H) [7], purchased from Boehringer, was used for the cleavage of 0.2 mg of S-carboxymethylated cinnamomin (1 mg/ml) with an enzyme to substrate weight ratio of 1:100, for 18 h at 37°C in 50 mM Na phosphate buffer, pH 8. Tryptic peptides were obtained using S-carboxymethylated protein, dissolved in 0.2 M N-ethylmorpholine acetate buffer, pH 8.2 (10 mg/ml), and incubated with trypsin (EC

Correspondence address: J.-C. Pernollet, Laboratoire d'Etude des Protéines, Centre INRA, route de St Cyr, 78026 Versailles Cedex, France

Abbreviations: Asp-N: Asp-N endoproteinase, EC 3.4.21; Cap, capsicein; Cin, cinnamomin; Cry, cryptogein; RPLC, reversed-phase liquid chromatography

Table 1
Edman amino acid sequencing of cinnamomin

Peptide denomination	N-term	T1	T2	T3	T4	T5	T6	T7	A	B
Elution position (CH ₃ CN%):	30	9	10	12	21	26	30	32.5	10	25
Cycle										
1	T (1630)	A(1580)	C(2970)	T(520)	L (710)	D(320)	T(350)	I (500)	D*(240)	I (147)
2	A(4300)	N(970)	A(3230)	A(1000)	M(890)	S (160)	L(630)	V(465)	G (170)	K(99)
3	C (2550)	G(670)	S (1740)	C(960)	C (590)	G (140)	V(490)	A(510)	F (180)	K(13)
4	T (3330)	F (630)	L(470)	T(340)	A (630)	Y (160)	S (580)	L(430)	S (28)	I (140)
5	A(3520)	S (560)		A(960)	S (530)	S (130)	I (470)	N(410)	S (10)	V(128)
6	T (2090)	S (-150)		T(310)	T (170)	M(81)	L(440)	P(420)	K (89)	A(140)
7	Q(2570)	K(180)		Q(800)	A (390)	L (80)	S (440)	P (-32)	C (100)	L(128)
8	Q(440)			Q(-17)	C (290)	T (28)	E(350)	D(170)	A (100)	N(134)
9	T (1890)			T(250)	N (290)	A (45)	S (150)	C(200)	S (20)	P(122)
10	A(2350)			A(590)	T (110)	T (25)	S (73)	D(99)	L (8)	P(8)
11	A(610)			A(26)	M(150)	A (28)	F (170)	L(190)		D(78)
12	Y (1810)			Y(390)	I (94)	L (50)	S (21)	T(68)		C(70)
13	K (1760)			K(230)	K (63)	P (42)	Q(130)	V(120)		D(51)
14	T (1180)					T (16)	C(67)	P(100)		L(41)
15	L (1270)					N (29)	S (40)	T(24)		T(33)
16	V (990)					A (20)	K(35)	S (110)		V(37)
17	S (1030)					□		G(51)		P(22)
18	I (810)							L(48)		T(16)
19	L (690)							V(43)		S(10)
20	S (560)							L(26)		G(12)
21	E (570)							D(27)		L(9)
22	S (260)							V(12)		V(5)
23	S (130)							Y(5)		L(7)
24	F (280)									D(11)
25	S (-43)									V(5)
26	Q(210)									Y(6)
27	C (190)									T(8)
28	S (40)									Y(1)
29	K (140)									A(2)
30	D(80)									N(3)
31	S (57)									G(2)
32	G (72)									F(3)
33	Y (84)									S (4)
34	S (25)									S (2)
35	M(43)									-
36	L (35)									C(1)
37	T (11)									A(1)
38	A(21)									
39	T (11)									
40	A(18)									
41	L (20)									
42	P (11)									
43	T (14)									
44	N(14)									
45	A (8)									
46	Q (9)									
47	Y (5)									
48	K (4)									
49	L (5)									
50	M(4)									
51	C (8)									
52	A (8)									

N-term, sequencing of the N-terminal end of the S-carboxymethylated cinnamomin. Values in parentheses indicate the differential amount (pmol) of phenylthiohydantoin amino acids determined at each step vs the former. Amino acids are indicated using the single-letter code. (□) Indicates that the sequencing process has been manually stopped, (*) that the corresponding residue has been deamidated. A, B and N-term were separated on a C8 Aquapore RP 300 BrownleeTM column, while tryptic peptides were chromatographed on a C8 ZorbaxTM column

3.4.21.4) treated with L-1-(tosylamido)-2-phenylethylchloromethylketone (enzyme to substrate weight ratio, 1:100) for 3 h at 37°C. The digestion was stopped by bringing the pH up to 2-3 with 6 N HCl. All digested materials were lyophilized prior to further separation.

Peptide fractionations by RPLC were conducted with an SP 8700 Spectra Physics equipment and monitored with a Jasco UV diode array detector. Tryptic peptides were separated at 40°C on a C8 ZorbaxTM 80 × 6.2 mm column at a flow rate of 1 ml · min⁻¹. A gradient

from 0 to 100% of solvent B was applied in 60 min; solvent A was MilliporeTM-treated water with 0.1% CF₃COOH and solvent B 40% MilliporeTM-treated water, 60% CH₃CN (FISONS, far UV grade) and 0.09% CF₃COOH. Asp-N peptides were obtained at room temperature with a C8 Aquapore RP 300 BrownleeTM 30 × 4.6 mm column at a flow rate of 0.5 ml · min⁻¹. A gradient from 0 to 100% of solvent B was applied in 60 min; in this case solvent B was 40% water, 60% CH₃CN and 0.09% CF₃COOH. CNBr peptides were separated with the same system; peptide B was rechromatographed at room temperature with a C8 Aquapore RP 300 BrownleeTM 30 × 2.1 mm column at a flow rate of 0.2 ml · min⁻¹. The following gradient was used: from 0 to 20% of solvent B in 5 min, then from 20% to 80% of solvent B in 55 min; solvent A was 5% CH₃CN in 25 mM CH₃COONH₄, pH 7.2, and solvent B 50% CH₃CN in 50 mM CH₃COONH₄, pH 7.2.

2.3. Protein sequencing and flexibility and hydropathy profile determination

Automated Edman degradation of the whole protein and of peptides was performed following the original method of Hewick et al. [8] using an Applied Biosystems 475A sequencer and its on line phenyl-thiohydantoin-amino acid analyzer model 120A with reagents and methods of the manufacturer. The hydropathy profiles were done with the programme of Kyte and Doolittle [9] and flexibility diagrams obtained with the method of Karplus and Schulz [10].

3. RESULTS

3.1. Isolation and sequencing of cinnamomin peptide fragments

Table 1 illustrates the sequencing data of cinnamomin and the derived peptides noted T (numbered according to their RPLC elution position), B and A for tryptic, CNBr and Asp-N cleavage products, respectively. With an initial sequencing yield of 45%, the S-

carboxymethylated N-terminal end was sequenced up to Ala52, which allowed the alignment of peptides T3, T6, T5 and T4. The sequencing of peptide B was sufficient to align the C-terminal end of the molecule (T4, T7, T1 and T2). Peptide A, obtained by partial cleavage at the amino side of the aspartic acid corresponding to a limited deamidation form of Asn89, confirmed the alignment of T₁ and T₂ and the C-terminus. Except that deamidation, no microheterogeneity was observed in the sequence of cinnamomin.

3.2. Amino acid sequence of cinnamomin

The alignment of the sequenced peptides of cinnamomin is shown in fig.1. Like cryptogin and capsicin, cinnamomin lacked some amino acids (Trp, His, Arg). It exhibited the same number of residues (98) and a similar *M_r* (10 282 instead of 10 323 for cryptogin and 10 155 for capsicin). Its calculated isoelectric point was 7.85, close to that of cryptogin because of the same number of lysyl residues (the calculated theoretical isoelectric points are 8.49 for cryptogin and 4.54 for capsicin).

3.3. Comparison of the structures of cinnamomin, cryptogin and capsicin

The 3 elicitors were homologous to each other and no deletion was necessary for aligning their sequences (fig.2). The percent match between cinnamomin and cryptogin was 87.8% and only 82.7% between cinnamomin and capsicin. The better homology with cryptogin was clearly observed in the N-terminal end

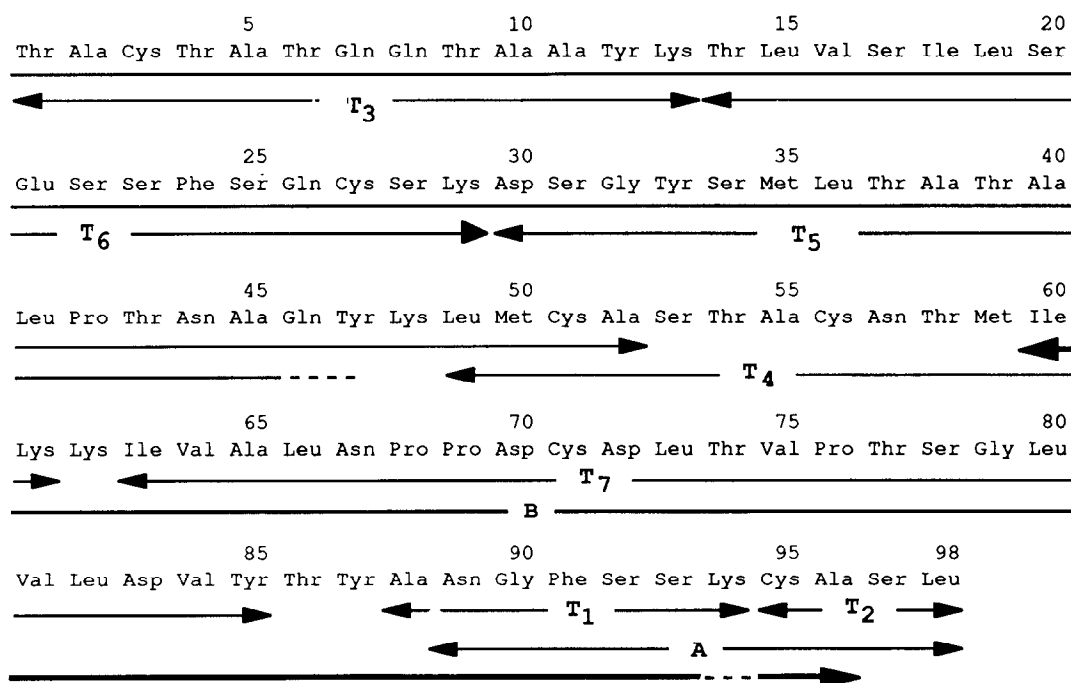


Fig.1. Complete amino acid sequence of cinnamomin secreted by *Phytophthora cinnamomi*. (→) N-terminus of the S-carboxymethylated protein; (↔) cleaved peptides (T, tryptic peptides; B, CNBr-cleaved peptides; A, endoprotease Asp-N-cleaved peptides). Dashed lines indicate undetermined residues.

					5						10					15					20
Cry	→	Thr	Ala	Cys	Thr	Ala	Thr	Gln	Gln	Thr	Ala	Ala	Tyr	Lys	Thr	Leu	Val	Ser	Ile	Leu	Ser
Cin	→	Thr	Ala	Cys	Thr	Ala	Thr	Gln	Gln	Thr	Ala	Ala	Tyr	Lys	Thr	Leu	Val	Ser	Ile	Leu	Ser
Cap	→	Ala	Thr	Cys	Thr	Thr	Thr	Gln	Gln	Thr	Ala	Ala	Tyr	Val	Ala	Leu	Val	Ser	Ile	Leu	Ser
					25						30					35					40
Cry	→	Asp	Ala	Ser	Phe	Asn	Gln	Cys	Ser	Thr	Asp	Ser	Gly	Tyr	Ser	Met	Leu	Thr	Ala	Lys	Ala
Cin	→	Glu	Ser	Ser	Phe	Ser	Gln	Cys	Ser	Lys	Asp	Ser	Gly	Tyr	Ser	Met	Leu	Thr	Ala	Thr	Ala
Cap	→	Asp	Ser	Ser	Phe	Asn	Gln	Cys	Ala	Thr	Asp	Ser	Gly	Tyr	Ser	Met	Leu	Thr	Ala	Thr	Ala
					45						50					55					60
Cry	→	Leu	Pro	Thr	Thr	Ala	Gln	Tyr	Lys	Leu	Met	Cys	Ala	Ser	Thr	Ala	Cys	Asn	Thr	Met	Ile
Cin	→	Leu	Pro	Thr	Asn	Ala	Gln	Tyr	Lys	Leu	Met	Cys	Ala	Ser	Thr	Ala	Cys	Asn	Thr	Met	Ile
Cap	→	Leu	Pro	Thr	Thr	Ala	Gln	Tyr	Lys	Leu	Met	Cys	Ala	Ser	Thr	Ala	Cys	Asn	Thr	Met	Ile
					65						70					75					80
Cry	→	Lys	Lys	Ile	Val	Thr	Leu	Asn	Pro	Pro	Asn	Cys	Asp	Leu	Thr	Val	Pro	Thr	Ser	Gly	Leu
Cin	→	Lys	Lys	Ile	Val	Ala	Leu	Asn	Pro	Pro	Asp	Cys	Asp	Leu	Thr	Val	Pro	Thr	Ser	Gly	Leu
Cap	→	Thr	Lys	Ile	Val	Ser	Leu	Asn	Pro	Pro	Asp	Cys	Glu	Leu	Thr	Val	Pro	Thr	Ser	Gly	Leu
					85						90					95					98
Cry	→	Val	Leu	Asn	Val	Tyr	Ser	Tyr	Ala	Asn	Gly	Phe	Ser	Asn	Lys	Cys	Ser	Ser	Leu		
Cin	→	Val	Leu	Asp	Val	Tyr	Thr	Tyr	Ala	Asn	Gly	Phe	Ser	Ser	Lys	Cys	Ala	Ser	Leu		
Cap	→	Val	Leu	Asn	Val	Tyr	Ser	Tyr	Ala	Asn	Gly	Phe	Ser	Ala	Thr	Cys	Ala	Ser	Leu		

Fig.2. Comparison of the amino acid sequences of cinnamomin, cryptogein and capsicein. Cry, cryptogein; Cin, cinnamomin; Cap, capsicein. Boxes emphasize differences.

up to residue 21. Most of the replacements were conservative [11]. With respect to cryptogein, 4 mutations were identical to those found in capsicein, for instance at positions 22 (Ala replaced by Ser), 39 (Lys replaced

by Thr), 70 (Asn replaced by Asp) and 96 (Ser replaced by Ala), while the 8 others are peculiar to cinnamomin (at positions 21, 25, 29, 44, 65, 83, 86 and 92). Like in cryptogein and capsicein sequences [5], internal se-

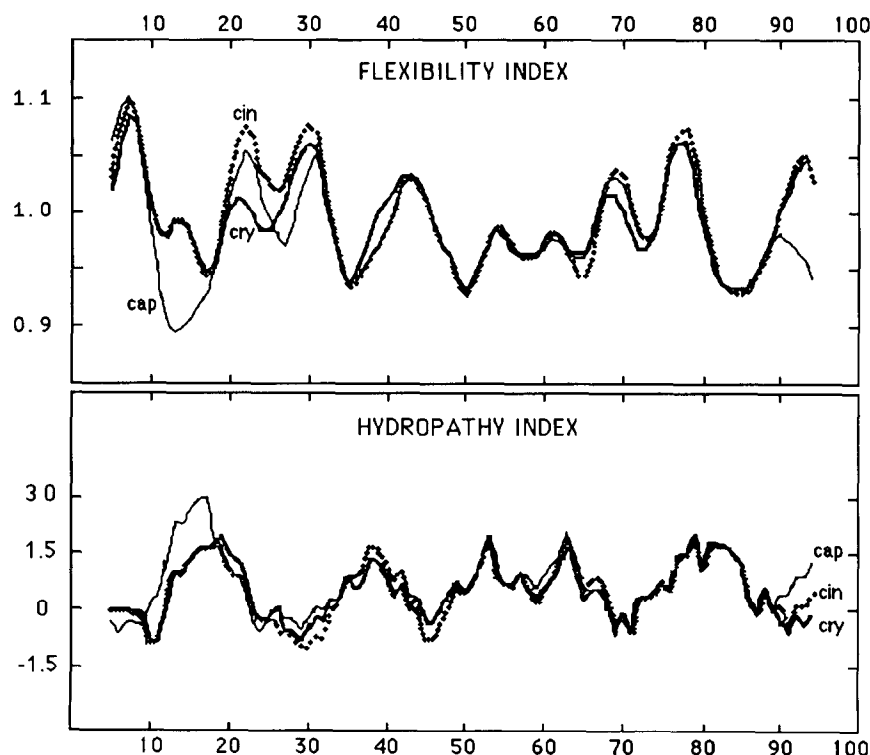


Fig.3. Buried and exposed areas of cinnamomin, cryptogein and capsicein. Profiles were obtained with a window width of 7. Flexibility diagrams were obtained with the method of Karplus and Schulz [10] and hydrophathy profiles computed with the programme of Kyte and Doolittle [9]. Bold lines correspond to cryptogein (cry), dashed lines to cinnamomin (cin), the others to capsicein (cap). All computations were done using the similarity matrix of Dayhoff et al. [11].

quence homologies consisting of the irregular repetition of 2 common consensus sequences of 6 to 7 residues occurred in cinnamomin.

The 3 flexibility and hydropathy profiles (fig.3) were very similar to each other, except at the C-terminal end and around position 15 where differences were prominent, while the 20–30 region was variable. In the 2 terminal regions, only capsicein showed differences, while cinnamomin and cryptogein profiles were very close together. Cinnamomin, with an average hydropathy index of 0.18 was comparable to cryptogein (0.13) but more hydrophilic than capsicein (0.39). The flexibility profiles reveal 6–7 putative antigenic determinants; only the C-terminal one was not predicted in capsicein and could be expected to differentiate it from the 2 other elicitors. The comparison of these diagrams with the primary structure showed that only 2 regions were to be involved in the difference of biological activity found between these molecules: an N-terminal region around position 13 where Lys in cinnamomin and cryptogein was replaced by Val in capsicein and a second one at the C-terminal end around position 94 involving the change of Lys in cinnamomin and cryptogein by Thr in capsicein.

4. DISCUSSION

With the exception of Asn89 which was partly deamidated, likely in the course of sample preparation, the absence of microheterogeneity all along the cinnamomin sequence establishes that cinnamomin, like other elicitors [5], arises from a single gene. Not only the amino acid sequences of cinnamomin and cryptogein were more homologous to each other than to capsicein, but they exhibit the same levels of biological activities [1,3,4]. Since the central core of the 3 molecules was highly conserved, it was likely essential for both biological activities. The terminal ends which vary from one elicitor to another would modulate differently these activities. Among the 15 replacements observed between cryptogein and capsicein, only those occurring at positions 13, 39, 61, 93 and 94 were not conservative. If one were to correlate the amino acid replacements with the biological activities, it is likely that neither the 5 replacements common to cinnamomin and capsicein nor those which were peculiar to cinnamomin were essential to explain the differences in biological activities, namely a reduction in remote leaf necrosis and an enhancement in the induction of the protection against *Phytophthora nicotianae* [5]. The only potential candidates were found at 9 positions: 1 (Thr replaced by Ala), 2 and 5 (Ala by Thr), 13 (Lys by Val), 14 (Thr by Ala), 28 (Ser by Ala), 61 (Lys by Thr), 72 (Asp by Glu) and 94 (Lys by Thr). The inversion between positions 1

and 2 and the mutations at positions 5 and 14, as well as those occurring at positions 28 and 72, were conservative [11]. Consequently, the replacements which are likely involved in the control of the level of necrotic activity with respect to the protective role of elicitors were restricted to only 3 positions (13, 61 and 94) instead of 5, as previously published [5]. In every case, a lysyl residue was involved.

When adding the comparison of the hydropathy plots, the role of Lys61 (which is next to another lysyl residue) appeared to be of lesser importance than the 2 strategic residues 13 and 94. As leaf symptoms occur at a distance from the application point and involve a systemic elicitor migration, differences in necrotic activity might also relate to features affecting the migration properties: not only are cinnamomin and cryptogein more hydrophilic than capsicein, but also their positive net charge would favour the negatively charged cell-wall crossing.

These results confirm more precisely those already inferred from the comparison of the sequences and deduced features of cryptogein and capsicein [5]. This information will be essential to engineer cryptogein-like molecules devoid of their necrotic properties to synthesize protective biochemicals against *Phytophthora*.

Acknowledgements: This study was supported by an INRA Action Incitative Programmée no. 89/4697. The authors wish to thank Dr Maud Bruneteau (Laboratoire de Biochimie microbienne, Université Claude Bernard-Lyon I) for providing cinnamomin and Monique Mansion for her skilful technical assistance.

REFERENCES

- [1] Bonnet, P. (1985) *Agronomie* 5, 801–808.
- [2] Bonnet, P., Poupet, A. and Bruneteau, M. (1985) *Agronomie* 5, 275–282.
- [3] Billard, V., Bruneteau, M., Bonnet, P., Ricci, P., Pernollet, J.C., Huet, J.C., Vergne, A., Richard, G. and Michel, G. (1988) *J. Chromatogr.* 44, 87–94.
- [4] Bonnet, P., Poupet, A., Abad, P., Venard, P. and Cardin, L. (1986) *Agronomie* 6, 829.
- [5] Ricci, P., Bonnet, P., Huet, J.C., Sallantin, M., Beauvais-Cante, F., Bruneteau, M., Billard, V., Michel, G. and Pernollet, J.C. (1989) *Eur. J. Biochem.*, in press.
- [6] Fernandez-Luna, J.L., Lopez-Otin, C., Soriano, F. and Mendez, E. (1985) *Biochemistry* 24, 861–867.
- [7] Ponstingl, H., Maier, G., Little, M. and Kraus, E. (1986) in: *Advanced Methods in Protein Microsequence Analysis* (Wittmann-Liebold, B. et al. eds) pp. 316–319, Springer, Berlin.
- [8] Hewick, R.M., Hunkapiller, M.W., Hood, L.E. and Dreyer, W.J. (1981) *J. Biol. Chem.* 256, 7990–7997.
- [9] Kyte, J. and Doolittle, R.F. (1982) *J. Mol. Biol.* 157, 105–132.
- [10] Karplus, P.A. and Schulz, G.E. (1985) *Naturwissenschaften* 72, 212–213.
- [11] Dayhoff, M.O., Schwartz, R.M. and Orcutt, B.C. (1978) in: *Atlas of Protein Sequence and Structure*, vol. 5, suppl. 3, pp. 345–352, National Biomedical Research Foundation, Washington, DC.